European Cytokine Network

European Cytokine Network. Vol. 11, Issue 3, September 2000: 362-71, Review

Three novel mammalian toll-like receptors: gene structure, expression, and evolution

Xin Du[1], **Alexander Poltorak**[1], **Yongie Wei**[2] **and Bruce Beutler**[1]

1. Scripps Research Institute, 10550 N. Torrey Pines Road, IMM-31, R204, La Jolla, CA 92037, USA.
2. Dept. of Physiology, The University of Texas, Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75234, USA.

Tirés à part
Reprints

Bruce Beutler, Scripps Research Institute, 10550 N. Torrey Pines Road, IMM-31, R204, La Jolla, CA 92037, USA. Tel: + 858 784 8609, Fax: + 1 858 784 8444, e-mail: bruce@scripps.edu

| RESUME | SUMMARY | ARTICLE, Part. 1, Part. 2, Part. 3, Part. 4 | REFERENCES | | FIGURES |

RESUME / SUMMARY

Haut de page

8/13/02

Mots clés

We describe three novel genes, encoding members of the Toll-like receptor (Tlr) family (TLR7, TLR8, and TLR9). These Tlr family members, unlike others reported to date, were identified within a genomic database. TLR7 and TLR8 each have three exons, two of which have coding function, and lie in close proximity to one another at Xp22, alongside a pseudogene. The remaining gene (TLR9) resides at 3p21.3 (in linkage with the MyD88 gene), and is expressed in at least two splice forms, one of which is monoexonic and one of which is biexonic, the latter encoding a protein with 57 additional amino acids at the N-terminus. The novel Tlrs comprise a cluster as nearest phylogenetic neighbors. Combining all sequence data related to Toll-like receptors, we have drawn several inferences concerning the phylogeny of vertebrate and invertebrate Tlrs. According to our best estimates, mammalian TLRs 1 and 6 diverged from a common mammalian ancestral gene 95 million years ago. TLR4, which encodes the endotoxin sensor in present-day mammals, emerged as a distinct entity 180 million years ago. TLRs 3 and 5 diverged from a common ancestral gene approximately 150 million years ago, as did Tlr7 and Tlr8. Very likely, fewer Tlrs existed during early vertebrate evolution: at most three or four were transmitted with the primordial vertebrate line. Phylogenetic data that we have adduced in the course of this work also suggest the existence of a *Drosophila* equivalent of MyD88, and indicate that the plasma membrane protein SIGIRR is close functional relative of MyD88 in mammals. Finally, a single present-day representative of the Toll-like proteins in *Drosophila* has striking cytoplasmic domain homology to mammalian Tlrs within the cluster that embraces TLRs 1, 2, 4, and 6. This would suggest that an ancestral (pre-vertebrate) Tlr may have adopted a pro-inflammatory function 500 million years ago.

Key-words Toll-like receptor, structure, phylogeny, chromosomal location, leucine-rich repeat, evolution, IL-1, IL-18, MyD88, SIGIRR.

ARTICLE

Haut de page

INTRODUCTION

The prototypic Toll protein of *Drosophila melanogaster* is a plasma membrane receptor characterized by a single transmembrane domain and a series of leucine-rich ectodomain repeats. It has been shown to fulfill both developmental and immunological functions, triggering dorsoventral patterning of the embryo [1, 2] and defending against fungal infections in adult life [3]. Similarly, the Toll ortholog 18-wheeler is required for defense against bacterial infections [4].

Mammals express an array of receptors with structural similarity to Toll. The first of these to be recognized was the IL-1 receptor [5, 6]. Subsequently, the IL-1 receptor accessory protein (IL-1RAcP), and both chains of the IL-18 receptor [7], were noted to have cytoplasmic domains with homology to the cytoplasmic domain of Toll. However, the ectodomain structures of these receptors (like that of the more recently discovered orphan receptor SIGIRR [8]) are based on repeats derived from the immunoglobulin superfamily, and are not leucine rich.

Quite apart from these cytokine receptors, a total of six orphan receptors with leucine-rich ectodomains and cytoplasmic Toll-like domains (*i.e.*, displaying homology to Toll over the full length of the protein rather than in the cytoplasmic domain alone) were identified by analysis of expressed sequence tag (EST) libraries [9-12]. Full-length (or nearly full-length) sequences were deduced for cDNAs encoding human Tlr1, Tlr2, Tlr3, Tlr4 [9, 11], and Tlr6 [12]; a partial sequence was determined for Tlr5 [11]. One of these receptors (Tlr4, or "h-Toll") was shown to be capable of triggering nuclear translocation of NF-kappaB upon ligation, in experiments in which chimeric receptors were expressed in mammalian cells, using a CD4 ectodomain in place of the native Tlr4 ectodomain [9].

The function of Tlr4 became clear with the recent positional cloning of the *Lps* locus, which revealed destructive mutations of *Tlr4* in endotoxin resistant mice of the C3H/HeJ and C57BL/10ScCr strains [13, 14]. In the former strain, a point mutation modifies the cytoplasmic domain of the Tlr4 protein, abrogating signal transduction. In the latter strain, the gene is deleted entirely by a mutation spanning 74,723 bp of genomic sequence [15]. Moreover, knockout of the *Tlr4* locus was shown to produce a phenocopy of the natural *Tlr4* mutations [16]. Hence, Tlr4 is

an essential component of the LPS signal transduction complex.

In further genetic complementation studies [17], the species origin of Tlr4 was shown to determine reactivity toward LPS partial structures that are known to be species-specific in their ability to stimulate macrophages. Hence, tetraacyl lipid A, an agonist for mouse macrophages, but an antagonist of LPS in human monocytic cell lines, was shown to be capable of activating mouse macrophages that expressed the mouse isoform of Tlr4, but was incapable of activating cells that expressed the human isoform of Tlr4. It would therefore seem certain that LPS is directly engaged by Tlr4, and does not act through an intermediary ligand.

Other microbial ligands, such as lipoteichoic acid [18], may also utilize the Tlr4 receptor. It would therefore appear that Tlr4 may be oligospecific in its ligand preference rather than monospecific. The *Tlr2* gene product was subsequently implicated in the transduction of signals evoked by Gram-positive bacteria [19], and knockout work has provided evidence that signaling initiated by peptidoglycan (muramyldipeptide) and at least some lipopeptides are both dependent upon this receptor [18, 20]. Hence Tlr2, like Tlr4, may be oligospecific rather than monospecific. It is plausible to think that other bacterial products signal *via* other Tlrs, and that in this manner, the Tlrs may collectively be responsible for the sensing of virtually all microbial pathogens.

For this reason, there has been a compelling motive for identifying novel Tlrs, though prospects for finding them among ESTs have diminished with the passage of time: at present, no strong homology may be found between any of the six known Tlrs and novel targets in dbEST, using the TBLASTN algorithm (unpublished observation). However, as genomic sequencing approaches a point of completion, it may be considered that all members of the Tlr superfamily will be identified.

As of this writing, 79.1% of the human genome has been published in the form of draft sequence, and 18.6% has been published as "finished" sequence. We therefore undertook to examine the human genomic sequence database in an effort to identify novel TLRs. We now detail the sequence, genomic structure, and expression characteristics of three novel TLRs (designated TLRs 7, 8, and 9 in accordance with the established convention) identified by this search. Two of these TLRs are X-linked, and lie in close apposition to one another at Xp22. A third is located at human chromosome 3p21.3. Though they are each represented as ESTs in one species or another, the extent of sequence overlap was insufficient to permit their recognition as "Toll-like" through a search of dbEST. The novel TLRs are longer than any of the existing TLRs, and comprise a new phylogenetic cluster.

MATERIALS AND METHODS

Search of genomic sequence for novel Tlrs

The Web-based NCBI BLAST program was used to examine three databases (Human Genome BLAST: finished human genomic sequence; the high-throughput genome sequence [htgs] database of draft sequence; and the nonredundant [nr] division of Genbank). The queries used were a region of the human MyD88 protein sequence corresponding to the area of homology with human Tlr4 (MyD88 residues 181-316), and also, the full-length human Tlr4 protein sequence. The TBLASTN algorithm (which translates the target sequence in all six frames before comparing with a peptide query) was used for the search. The e-value of the output was set at 10.

PCR amplification of the target human genomic sequence

Human placental cDNA purchased from Clontech was used as the template for amplification of Tlr7 and Tlr8 sequences. THP-1 cell mRNA was used as a template for the reverse transcription and amplification of Tlr9 cDNA sequences. Initial amplification of TLR7 was accomplished using the primers:

AGT GGA AAT TGC CCT CGT TGT TAT AAT GCC
and
CCA CTC GGT CAC AGC TGG GTC TTT AGT G
Initial amplification of TLR8 was accomplished using the primers:

AAC ATA GAA GAT GGA GTA TTT GAA ACG CTG AC
and
TCA AAT ACT GAG AAT GCT GTA ACA CTG GCT C
Initial amplification of TLR9 was accomplished using the primers:
GCA TGC CCT GCG CTT CCT ATT CAT GGA C
and
CCA TAG ACC GAG GCC CAC AGG TTC TCA AAG

cDNA cloning of three novel human Tlrs

THP-1 cell RNA was reverse transcribed using an oligo dT primer. The same primers used for initial genomic amplifications were used to amplify the major part of the corresponding cDNAs. The 5' and 3' termini of the mRNA molecules were determined by 5' and 3' RACE. For these reactions, the following primers were used:

TLR7: 5' RACE- TGG GGG CAC ATG CTG AAG AGA GTT A

TLR7: 3' RACE- AGG CCA AGA TAA AGG GGT ATC AGC GTC T

TLR8: 5' RACE- ATG CCC CAG AGG CTA TTT CTC CCA C

TLR8: 3' RACE- CTT TAT TAT GGC GCG AAA TCA TGA CT

TLR9: 5' RACE- GCA TCA GGA TGT TGG TAT GGC TGA GGG A

TLR9: 3' RACE- AAG GAG CTG CGA GAG CTC AAC CTT AGC G

DNA sequencing

Using the primers shown in Table 1, the entire cDNA of each of the novel Tlrs, and the coding region of the genomic sequences of each of the novel TLRs, was determined using Model 373 and 377 DNA sequencers, and amplified templates. Dye-terminator chemistry was used, and trace files were assembled with the programs phred and Phrap (obtained from the University of Washington Genome Center). The alignment of genomic and cDNA sequences was performed using the fasta program of the GCG 9.0 suite, and mismatches were examined using the program Consed (University of Washington Genome Center).

Bioinformatic analysis

As already mentioned, proteins exhibiting Toll-like cytoplasmic domains were identified by defining the most conserved region of the Toll-like domain in human MyD88, a rather distant member of the family. Residues 181-316 of human MyD88 were then designated as a query sequence to be used in finding all homologous sequences by BLAST search, carried out against the human genome (using TBLASTN) and against the nonredundant Genbank database (using BLASTP). The GCG program Pileup was then used to align the polypeptide sequences of Tlrs 1 though 9 and all other proteins or open reading frames known to have homology to the Toll-like domain of human MyD88. A gap penalty setting of 4 and a gap extension penalty of 2 were assigned for this purpose.

Evolutionary distance assessments were performed using the program PAUP (Smithsonian Institution), run with the Seqlab modality of GCG, using the setting for maximum parsimony. A bootstrapping procedure was followed, with 100 repetitive measurements performed to determine tree structure. Only a single tree was generated for each set of input data. Final trees were drawn by hand, according to output data generated by PAUP. In dating divergence, an neutral drift model was used, and all distances were calculated relative to the average separation of human and mouse TLR2, TLR4, TLR5, TLR6, and TLR7 genes. Measurements were confirmed by determination of synonymous (third codon) substituiton rates for all coding sequences. All numbers presented were calculated based on the tenable assumption that humans and mice diverged from a common ancestor 60 million years ago.

The program Generunner was used for analysis of reading frames. The program SMART (http://smart.emblheidelberg.de/) was used for comparative analysis of structural domains.

RESULTS

Six Tlr cDNAs were known to be present in Genbank prior to the initiation of this study. All of these (Tlrs 1, 2, 3, 4, 5, and 6) were represented in the high-throughput genomic sequencing databases (htgs; databases encompassing phase 1 and phase 3 sequencing efforts). Three additional high-scoring homologies were identified on the human X chromosome (Xp22.2 to 22.3), and one high-scoring homology was identified on chromosome 3 (3p21.3). The triad of X-chromosome homologs are distributed over approximately 55 kb of genomic DNA, and lie in close proximity to a pair of PRPS2 (phosphoribosylpyrophosphate synthetase II) genes; the single chromosome 3 homolog lies in close proximity to ALAS1 (the delta-aminolevulinate synthase gene), and is also in the same cytogenetic interval as the MyD88 gene.

Two of the X-chromosome homologs and the chromosome 3 homolog had long open reading frames; the remaining X-chromosome homolog, lying in opposite orientation to the first two and placed to one side of them, did not. Reference to the dbEST database immediately revealed that the two X-chromosome homologs with long open reading frames were expressed as processed mRNA, in that each had several EST representatives. The chromosome 3 homolog was not initially represented by a human transcript in dbEST, but one such EST was added in the course of our work (accession no. AW502545). The remaining X-chromosome homolog was not represented at all, and for this reason as well as the fact that it lacked a long open reading frame, it was presumed to be a pseudogene.

Each Tlr cDNA was amplified by PCR and sequenced in entirety. The furthest 5' extent of each mRNA was determined by RACE. On this basis, the presumed full-length sequence of the two expressed X-chromosomal Tlr cDNAs and the chromosome 3 Tlr cDNAs were submitted to Genbank as Tlrs 7, 8, and 9, respectively (AF240467 [TLR7], AF246971 [TLR8], and AF259262 and AF259263 [TLR9]). In Figure 1, the genomic structure of each of the novel TLRs is presented, as deduced from the cDNA sequence. Both TLR7 and TLR8 have two introns and three exons, much as described for *Tlr2* [18] and *Tlr4* [14, 21]. However, as distinct from both *Tlr2* and *Tlr4*, only exon 2 and exon 3 have coding function in TLR7 and TLR8. Moreover, only the initiator methionine is encoded by exon 2 in either gene. Hence, the coding region of TLR7 and TLR8 genes is almost entirely monoexonic. Two separate splice forms of TLR9 were cloned by RACE. One had two exons, while the other had a single exon. The biexonic protein (designated Tlr9a) has 57 additional amino acids at the N-terminus as compared with the monoexonic protein (Tlr9b) (Figure 2).

On the basis of these findings, the length of each new gene can be assigned. From cap site to poly-A addition site, the TLR7 gene is 23,279 bp in length; the TLR8 gene is 15,563 bp in length; the TLR9 gene is either 5,069 or 3,109 bp in length, depending upon definition. TLR7 and TLR8 are separated from each other by 16,273 bp of intervening DNA and are transcribed in the same direction, though the polarity of transcription with respect to the centromere remains unknown.

The inferred structure of the proteins specified by these three novel TLRs are interesting in that each is longer than any of the Tlrs reported previously (see alignment, Figure 2). Moreover, there are more leucine-rich repeats than are present in most of the other Tlrs (15 for TLR7; 16 for TLR8; 18 for TLR9). Only Tlr3, which has 19 repeats, falls among the new Tlrs in this respect. All of the novel TLRs exhibit a proline residue at the site corresponding the mutation that renders C3H/HeJ mice unresponsive to LPS; TLR3 alone provides the exception to the rule. Some residues are even more stringently conserved, and are in strict consensus throughout the family. mRNA encoding each of the novel Tlrs is expressed at a very low level in mammalian tissues. As such, only a weak hybridization signal could be obtained using any of the full-length cDNA molecules as probes on Northern blots (not shown). However, a sense of the expression pattern of these novel genes was obtained using RT-PCR (Figure 3).

The evolutionary relationship among the nine existing mammalian Toll-like receptors was examined using the method of maximum parsimony after performing an optimal alignment using Pileup. A single solution was obtained, regardless of parameter variation; in particular, whether or not bootstrapping was used. In an unrooted tree denoting the these nine mammalian receptors alone, TLR7, TLR8, and TLR9 emerge as nearest neighbors (Figure 4). TLR1 and TLR6 are the most closely related TLRs of the nine examined; they, in turn were most nearly related to TLR2, and then, to TLR4. TLR3 and TLR5 appear to have diverged from a common

ancestral gene, and both were isolated from all other mammalian Tlrs by a considerable distance. An informed estimate of the dates of divergence is presented in Table II.

The relationship of the mammalian Tlrs with all other proteins bearing a Toll-like cytoplasmic domain was then examined. This independent approach to phylogenetic assignment (Figure 5) disclosed a relationship among the mammalian TLRs identical to that obtained using whole protein sequences, and barely distorted at all by exclusion of the ectodomain. As with the former method - and more obviously so - cytoplasmic domain homology relegates TLR7, TLR8, and TLR9 to a cluster of their own. TLR5 and TLR3 are still more dramatically isolated from the other TLRs. TLR1 and TLR6 are seen to be more similar to one another than any of the other TLRs, and next most closely related to TLR2, and then, to TLR4.

Some surprises are evident. A *Drosophila* protein (cg5528) is found to even more similar to the [TLR1, TLR6, TLR2] cluster than is TLR4. Though one of the nine *Drosophila* Toll-like proteins now known to be endowed with a leucine-rich ectodomain, cg5528 was grouped with the mammalian proteins solely on the basis of cytoplasmic domain homology. Another *Drosophila* protein (cg2078) also appears within a mammalian cluster: this time, in close proximity to MyD88 and SIGIRR. The IL-1 and IL-18 receptors, and the related ST2L protein lie within a distinct cluster of their own.

DISCUSSION

The identification of three novel TLR genes (here designated TLR7, TLR8, and TLR9) was a direct consequence of high-throughput genomic sequencing. Though retrospectively it is possible to see that ESTs derived from each of these genes were represented in dbEST, all of these ESTs were either too short or improperly placed, so as to preclude significant matches with the defining Toll-like domain of any of the existing TLRs, using either BLASTN or TBLASTN search algorithms. Though the ESTs might have served to identify the anonymous genomic seqences of these TLRs as genes, the inevitable fact of EST database contamination by genomic DNA makes EST identification a somewhat risky standalone criterion to use in gene identification.

The total number of TLRs yet to be found in the mammalian genome remains a matter of speculation, although by this time, it is clear that at least the majority of them have been found, and very likely all of them. All six of the earlier-described Toll-like receptor molecules are among these. It is therefore reasonable to discuss the phylogenetic relationships that unite these TLRs with each other, and with other proteins that bear Toll-like cytoplasmic domains.

In the course of this work, a large number of genomic sequences bearing similarity to the ectodomain-encoding region of the true Tlrs were identified. Among these sequences, which probably bear the leucine-rich repeat module that typifies Tlrs, it is quite likely that some authentic genes are to be found. However, the function of proteins with leucine-rich motifs is quite variable, and in the main, is likely to be unrelated to the function of the Tlrs. As such, we made no immediate attempt to identify or characterize the expression patterns of genes that were bereft of a Toll-like domain.

The leucine-rich repeat has been utilized for many purposes in evolution. It is rather poorly conserved, and several proteins that lack Toll-like cytoplasmic domains - indeed, some secreted proteins as well as surface receptors - are recognized by BLAST searches based on the ectodomain sequence of Toll-like receptor ectodomains. By contrast, Toll-like homology represented in the cytoplasmic domain of all nine mammalian Tlrs, the adapter protein MyD88, the *Drosophila* Toll-like receptors, members of the IL-1 receptor family, and certain orphan receptors (for example, SIGIRR, and ST2L) is far more conserved; *i.e.*, it comprises a slower molecular clock with which to assess evolutionary relationships. As previously noted, the Tlr4 ectodomain is far more variable among species that the Tlr4 cytoplasmic domain [21]. Further, the common structure of the cytoplasmic domain of the true Toll-like receptors bespeaks an overlap of function, already apparent from what is known of some of these receptors. For example, the IL-1, IL-18, Tlr4, and Tlr2 receptors all depend upon MyD88 for signal transduction, and all are known to deliver pro-inflammatory signals. From all species, 47 molecules bearing Toll-like cytoplasmic domains may now be assembled. It is, therefore, possible to answer a number of questions concerning functional and phylogenetic relationships.

It is to be noted that divergence implies an increase in the number of genes of a given type.

While some genes may become non-functional (*i.e.*, pseudogenes) and degenerate to an unrecognizable state under the pressure of mutation, several hundred million years are required for such "resorption" to occur [22]. At present, nine Tlr genes and one Tlr pseudogene are known to reside in the human genome. However, as may be deduced from Table 2, 95 million years ago, there were but nine genes at most, and just over 150 million years ago, there were but six genes. Our best estimate holds that the progenitor of *Drosophila* and humans bequeathed at most four genes bearing Toll-like domains to the descendant phyla. In *Drosophila* and humans alike, nine true Toll-like receptors may be found today.

Where TLR4 is concerned, a related inference immediately presents itself. Mammals and birds are known to be LPS-sensitive (the latter most obviously so during embryonic life [23]), and while evidence exists for a reptilian response [24-30], reptiles almost surely less sensitive. No other class in the subphylum *Vertebrata* is known to respond to LPS. Mammals, reptiles and birds are believed to have evolved from a common ancestor no more recently than 350 million years ago. It may be supposed that each line inherited the primordial gene that became mammalian TLR4 from this ancestor. In each descendant line, and particularly in mammals and birds, a mechanism for LPS sensing has evolved, whereas LPS sensing seemingly has not evolved in other vertebrate lines [31]. The development of LPS sensitivity seems, then, to have been an example of convergent evolution in mammals and birds, perhaps developed as a result of changes in microbial flora that attended the development of a warm-blooded way of life. However, it is somewhat problematic to speak of an avian "Tlr4," since Tlr4 did not exist as a separate entity until long after mammals and birds diverged from one another.

Several additional facts are apparent from the phylogenetic tree presented in Figure 5. First, since *Drosophila* proteins are grouped with mammalian proteins on the basis of cytoplasmic domain similarity, it may be inferred that some of the signaling functions present in modern-day fruit flies and mammals were already present in a common ancestor, some five hundred million years ago. It may be speculated that cg5528, which bears a cytoplasmic domain more similar to TLR1, TLR6 and TLR2 than any of the other mammalian TLRs, transmits a pro-inflammatory signal similar to that assigned to TLR2 or TLR4. The commonality of structure within the cytoplasmic domain of these molecules bespeaks an ancestry more ancient than might previously have been supposed.

Further, though no mutational evidence has yet emerged in support of the possibility, it can reasonably be proposed that cg2078 - found in a cluster with SIGIRR and MyD88 - acts in *Drosophila* to transduce signals much as MyD88 does in mammals. Further to this view, the cg2078 protein has no leucine-rich ectodomain, and indeed, appears to be strictly cytoplasmic: a fact quite independent of its assignment to the MyD88/SIGIRR cluster. We also suggest that, in mammals, SIGIRR may fulfill a role in signaling quite similar to that of MyD88, this though SIGIRR is a transmembrane protein whereas MyD88 is not.

Among *Drosophila* transmembrane proteins with Toll-like cytoplasmic domains, there are no representatives with ectodomains that display immunoglobulin-type repeats. This might be taken as evidence that the IL-1R/IL-18R/ST2L cluster arose after invertebrates and vertebrates diverged. Indeed, the relatively tight grouping of this cluster supports the assertion of a recent evolutionary origin. Moreover, insofar as IL-1R and IL-18R are linked to the adaptive immune response rather than to primary detection of microbes by the innate immune system, a more recent phylogenetic appearance would be expected.

The fact that two X-chromosomal TLRs lie in close approximation to one another (alongside a presumed TLR pseudogene), and the added fact that TLR9 lies in the same cytogenetic interval as the MyD88 (a cytoplasmic transducer with a Toll-like domain) suggests that at least some members of the family arose from the tandem duplication of ancestral genes. Concordant with the hypothesis of a recent duplication event, the two X-chromosome representatives are quite closely related to one another. However, they are less structurally similar than TLRs 1 and 6, which reside on different chromosomes, but likely deliver very similar signals.

The assignment of function to distinct members of the Tlr family is of paramount importance in this rapidly developing field. We are able to draw no substantial inferences based on the distribution of TLR gene expression. The novel TLRs, like most TLRs, seem broadly expressed, though differentially expressed, and expressed at very low levels compared to most "housekeeping" genes. So far, the best understood members of the Tlr family are certainly Tlr4

and Tlr2. The function of Tlr4 as an LPS receptor has been demonstrated by the identification of two naturally occurring mutations that render mice resistant to LPS [13, 14], and subsequently was confirmed by knockout of the gene [16]. It would appear that Tlr2 serves to identify lipopeptides [20] and peptidoglycan [18]. The function of all other members of the family remains unknown, though there is speculation that both immunologic functions (related to microbial sensing by the innate immune system) and developmental functions (akin to dorso-ventral patterning of the embryo subserved by *Drosophila* Toll) might be ascribed to individual Tlr proteins. It would seem that the identification or production of mutations in TLR genes comprise the best tools presently available for the assignment of function.

From this standpoint, it is particularly fortunate that two of the novel TLRs (TLR7 and TLR8) reside on the X chromosome. As such, phenotypically detectable mutations should be recoverable in males, and in some females, depending upon the extent of X-inactivation to which the normal allele is subject. Moreover, hemizygosity for knockout mutations of the X-chromosomal loci is conveniently achieved. Finally, insofar as the mammalian X chromosome is conserved *in toto*, it may be expected with confidence that TLR7 and TLR8 reside on the X chromosome in other mammalian species as in humans.

We are unaware of unexplained immunologic phenotypes associated with Xp22. However a number of developmental phenotypes map to the region and remain unexplained, and recently, mutations affecting a novel ortholog of the IL-1 receptor mapping to Xp22.3 have been associated with an X-linked form of mental retardation [32]. As to the 3p21.3 location of TLR9, we are also unable to point to phenotypes that seem certain to be caused by mutations of this gene, but note that several neoplastic diseases have been mapped to this region. The possibility that this Tlr, or other Tlrs yet to be identified, might act as tumor suppressors is a plausible one, since there is direct continuity between the Toll-like domain and pro-apoptotic pathways within the cell. MyD88 and IRAK, for example, both downstream intermediates for Tlr signal transduction, have well-defined death domains. This may account for some of the effects wrought by Tlr agonists: for example, the pro-apoptotic effect of LPS applied to cultured macrophages [33].

CONCLUSION

Haut de page

Acknowledgements. This work was supported by NIH grant no. RO1 GM60031-01A1. In the course of these studies, we became aware that Chuang and coworkers had independently isolated the same collection of novel Tlrs as we. By mutual agreement, their work has been reported back-to-back with our own. We are grateful to Dr. Anna DiRienzo for helpful discussion of evolutionary distances.

REFERENCES

Haut de page

1. Anderson K V, Jurgens G, Nusslein-Volhard C. 1985. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: genetic studies on the role of the Toll gene product. *Cell* 42: 779.

2. Anderson K V, Bokla L, Nusslein-Volhard C. 1985. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: the induction of polarity by the Toll gene product. *Cell* 42: 791.

3. Lemaitre B, Nicolas E, Michaut L, Reichhart J M, Hoffmann J A. 1996. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* 86: 973.

4. Williams M J, Rodriguez A, Kimbrell D A, Eldon E D. 1997. The 18-wheeler mutation reveals complex antibacterial gene regulation in *Drosophila* host defense. *EMBO J.* 16: 6120.

5. Gay N J, Keith F J. 1991. *Drosophila* Toll and IL-1 receptor [letter]. *Nature* 351: 355.

6. Heguy A, Baldari C T, Macchia G, Telford J L, Melli M. 1992. Amino acids conserved in interleukin-1 receptors (IL-1Rs) and the *Drosophila* toll protein are essential for IL-1R signal

transduction. *J. Biol. Chem.* 267: 2605.

7. Born T L, Thomassen E, Bird T A, Sims J E. 1998. Cloning of a novel receptor subunit, AcPL, required for interleukin-18 signaling. *J. Biol. Chem.* 273: 29445.

8. Thomassen E, Renshaw B R, Sims J E. 1999. Identification and characterization of SIGIRR, a molecule representing a novel subtype of the IL-1R superfamily. *Cytokine* 11: 389.

9. Medzhitov R, Preston-Hurlburt P, Janeway C A Jr. 1997. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature* 388: 394.

10. Chaudhary P M, Ferguson C, Nguyen V, Nguyen O, Massa H F, Eby M, Jasmin A, Trask B J, Hood L, Nelson P S. 1998. Cloning and characterization of two Toll/Interleukin-1 receptor-like genes TIL3 and TIL4: evidence for a multi-gene receptor family in humans. *Blood* 91: 4020.

11. Rock F L, Hardiman G, Timans J C, Kastelein R A, Bazan J F. 1998. A family of human receptors structurally related to *Drosophila* Toll. *Proc. Natl. Acad. Sci. USA* 95: 588.

12. Takeuchi O, Kawai T, Sanjo H, Copeland N G, Gilbert D J, Jenkins N A, Takeda K, Akira S. 1999. TLR6: a novel member of an expanding Toll-like receptor family. *Gene* 231: 59.

13. Poltorak A, Smirnova I, He X L, Liu M Y, Van Huffel C, McNally O, Birdwell D, Alejos E, Silva M, Du X, Thompson P, Chan E K L, Ledesma J, Roe B, Clifton S, Vogel S N, Beutler B. 1998. Genetic and physical mapping of the *Lps* locus- identification of the toll-4 receptor as a candidate gene in the critical region. *Blood Cells Mol. Dis.* 24: 340.

14. Poltorak A, He X, Smirnova I, Liu M Y, Van Huffel C, Du X, Birdwell D, Alejos E, Silva M, Galanos C, Freudenberg M, Ricciardi-Castagnoli P, Layton B, Beutler B. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in *Tlr4* gene. *Science* 282: 2085.

15. Poltorak A, Smirnova I, Clisch R, Beutler B. 2000. Limits of a deletion spanning *Tlr4* in C57BL/10ScCr mice. *Submitted.*

16. Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, Takeda K, Akira S. 1999. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *J. Immunol* 162: 3749.

17. Poltorak A, Ricciardi-Castagnoli P, Citterio A, Beutler B. 2000. Physical contact between LPS and Tlr4 revealed by genetic complementation. *Proc. Natl. Acad. Sci. USA* 97: 2163.

18. Takeuchi O, Hoshino K, Kawai T, Sanjo H, Takada H, Ogawa T, Takeda K, Akira S. 1999. Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. *Immunity* 11: 443.

19. Underhill D M, Ozinsky A, Hajjar A M, Stevens A, Wilson C B, Bassetti M, Aderem A. 1999. The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens. *Nature* 401: 811.

20. Takeuchi O, Kaufmann A, Grote K, Kawai T, Hoshino K, Morr M, Mühlradt P F, Akira S. 2000. Cutting edge: preferentially the R-stereoisomer of the mycoplasmal lipopeptide macrophage-activating lipopeptide-2 activates immune cells through a Toll-like receptor 2- and MyD88-dependent signaling pathway. *J. Immunol.* 164: 554.

21. Smirnova I, Poltorak A, Chan E K, McBride C, Beutler B. 2000. Phylogenetic variation and polymorphism at the Toll-like receptor 4 locus (TLR4). *Genome Biology* 1: 1.

22. Li W H. 1997. In: *Molecular Evolution*, 177-214. Sinauer Associates, Inc., Sunderland, MA.

23. Finkelstein R A. 1964. Observations on mode of action of endotoxin in chick embryos. *Proc. Soc. Exp. Biol. Med.* 115: 702.

24. Saad A H, el Deele S. 1990. Immunological changes during pregnancy in the viviparous

lizard, Chalcides ocellatus. *Vet. Immunol. Immunopathol.* 25: 279.

25. Saad A H. 1989. Sex-associated differences in the mitogenic responsiveness of snake blood lymphocytes. *Dev. Comp. Immunol.* 13: 225.

26. Clark I A. 1982. Correlation between susceptibility to malaria and babesia parasites and to endotoxicity. *Trans. R. Soc. Trop. Med. Hyg.* 76: 4.

27. Zurovsky Y, Brain T, Laburn H, Mitchell D. 1987. Pyrogens fail to produce fever in the snakes Psammophis phillipsii and Lamprophis fuliginosus. *Comp. Biochem. Physiol. A.* 87: 911.

28. Zurovsky Y, Mitchell D, Laburn H. 1987. Pyrogens fail to produce fever in the leopard tortoise Geochelone pardalis. *Comp. Biochem. Physiol. A.* 87: 467.

29. Saad A H, Shoukrey N. 1988. Sexual dimorphism on the immune responses of the snake, Psammophis sibilans. *Immunobiology* 177: 404.

30. McKinney E C, Bentley T B. 1985. Cell-mediated immune response of Chelonia mydas. *Dev. Comp. Immunol.* 9: 445.

31. Berczi I, Bertok L, Bereznai T. 1966. Comparative studies on the toxicity of *Escherichia coli* lipopolysaccharide endotoxin in various animal species. *Can. J. Microbiol.* 12: 1070.

32. Carrie A, Jun L, Bienvenu T, Vinet M C, McDonell N, Couvert P, Zemni R, Cardona A, Van Buggenhout G, Frints S, Hamel B, Moraine C, Ropers H H, Strom T, Howell G R, Whittaker A, Ross M T, Kahn A, Fryns J P, Beldjord C, Marynen P, Chelly J. 1999. A new member of the IL-1 receptor family highly expressed in hippocampus and involved in X-linked mental retardation. *Nat. Genet.* 23: 25.

33. Karahashi H, Amano F. 1999. LPS-induced signals in activation of caspase-3-like protease, a key enzyme regulating apoptotic cell damage into a macrophage-like cell line, J774.1, in the presence of cycloheximide. *J. Leukoc. Biol.* 66: 689.

Haut de page